University of Bern

Faculty of Humanities
Department of Linguistics and Literary Studies

Institute of Linguistics
Historical Comparative Linguistics

Prof. Dr. Roland Bielmeier

Bachelor thesis

July 31st 2009

66’338 characters

A Comparison of a Manuscript with the Indus Script

Lucy Zuberbühler
Jurastrasse 7
3063 Ittigen

lucy.zuberbuhler@students.unibe.ch
031 921 63 29 / 078 641 92 66
06-126-627

Table of contents

Table of figures

Introduction

For over 130 years scholars have been struggling to decipher the intriguing and perplexing system of symbols, which was used in and around the Indus Valley by one of the most ancient and extensive urban civilizations. Labouring under the almost universally held assumption that the discovered strings of signs encoded the language or languages of the Indus Culture, hundreds of attempts at reading the supposed script have been undertaken, but so far nobody has succeeded in convincingly translating the markings found on stamp seals and a dozen other media (Lawler 2004: 2026).

The fact that the sign sequences withstood every academic effort at decipherment, combined with the troubling lack of sign repetitions and the brevity of the inscriptions, recently prompted some scholars to assert that the Indus symbols might not have been intended to encode speech at all, but might instead be a system of religious, political and social symbols not tied to any particular language (Farmer / Sproat / Witzel 2004: 19). This proffered theory, which has gained a growing number of supporters of late, rests predominantly on the finding that no Indus inscriptions of considerable length have ever come to light (Lawler 2004: 2026). The longest discovered Indus text contains a mere 17 signs and the average inscribed object shows only 4-5 signs, which seems a suspiciously low amount considering the vast number of Indus artefacts that have been unearthed. This lack of evidence for a manuscript tradition is taken to be an indicator that the Indus Civilization was not in fact literate and never used a linguistic system of writing (Farmer / Sproat / Witzel 2004: 19-22).

This provocative hypothesis, which has riled Indian nationalists and caused vehement protests of dissent amongst academics working in the field of Indus studies, would, however, be considered falsified under certain circumstances (Lawler 2004: 2026). Namely, if an Indus text, containing several hundred signs, were ever discovered, or if an Indus inscription were found, which bore at least 50 symbols and additionally showed the sign repetitions expected in a script (Farmer / Sproat / Witzel 2004: 48). Discoveries of such a nature would not only confirm that the earliest Indian civilization employed a writing system, but might also assist with the breakthrough decipherment of one of the few remaining major undeciphered scripts. This, in turn, would shed light on

many unanswered questions concerning the fascinating Indus Civilization (Robinson 2007: 145).

An artefact, that could potentially fulfil these two requirements and would therefore be of paramount importance to the Indus studies, has surfaced in an Afghani museum collection. The artefact in question consists of a strip of what would appear to be several thin layers of bark, with seven lines of symbols running across it. A cursory perusal of the approximately 200 symbols, could lead a layperson with a rudimentary knowledge of the Indus script, to propose that the text at hand might have been composed using signs that appear in the Indus script, for some of the most common and well-known Indus signs seem to be present on the bark manuscript. It remains to be seen, however, exactly how many of the graphs on the bark object could be considered attested Indus symbols.

A detailed study of this exciting artefact and a subsequent comparison with the Indus script would therefore prove to be a worthwhile and necessary endeavour, which is why an analysis of this previously unexamined manuscript fragment will be conducted in this thesis. The aim here is to perform a graphemic analysis, segmenting the strings of markings into individual symbols and separating sign variants from distinct signs, thus allowing a sign list to be put forward and a sign repetition count to be tallied. In a next stage the proposed sign list will be compared to Indus symbols, with the goal of determining the number and extent of matching correspondences between the manuscript characters and the Indus symbols. Before a comparison of this nature can be carried out though, an overview of what is known about the Indus script will be required, so this essential background information will be presented first.

1. The Indus script

1.1. Development of the Indus Civilization and its script

Like the ancient empires of Egypt and Mesopotamia, the civilization, which created the Indus script, was able to prosper because it made excellent use of the resources that river valleys provide (Joshi and Parpola 1987: XI). In the case of the Indus Civilization, it was the vast Indus Valley in modern-day Pakistan, which attracted settlers because the surrounding mountainous areas provided a variety of useful raw materials, while its

vast plain offered abundant agricultural resources (Kenoyer 2006: 26-28). This great plain is exceptionally fertile due to the immense amount of alluvium that the Indus River deposits on its banks. The river possesses this silt displacement capacity because the melting Himalayan snow and the precipitation of the monsoon season feed it and its tributaries with an enormous quantity of water (Jarrige 1988: 13).

Small farming communities started to cultivate this fertile land in the Indus highland valleys from about the seventh millennium BC onwards. This early food-production era was followed in about 4’300 BC by a period of migratory expansion and trade network establishment. Village size increased and groups spread to the Sarasvati River, a parallel running, once mighty river system to the east, which could have supported the existence of a great number of people at that point in time (Parpola 2005: 29). Linking up major river systems and locating routes through mountain passes was vital, as it facilitated the transport of goods between costal, plain and mountain communities and ensured that the population had access to a host of resources, which could compensate for insufficient regional crop harvests (Kenoyer 2006: 26-30).

This prolonged contact between the dispersed villages also meant that a considerable cultural uniformity was developing after about 3’200 BC. The uniformity can be observed in pottery, jewellery and figurine similarities, in the common employment of particular types of stamp seals and potters’ marks, as well as in the layout of some of the settlements (Parpola 2005: 29-30). These shared characteristics of culture, which did vary somewhat from region to region, are regarded as the early stages of the civilization referred to as the Indus or Harappan Civilization (Possehl 1996: 10-11).

This great civilization matured and reached its developmental pinnacle in the period between approximately 2’500 and 1’900 BC. It was in the short transition period leading from the Early Harappan to the Mature Harappan period that the Indus script emerged. It did so in its fully developed state, showing no signs of having passed through a formative phase, although some of the symbols did already appear as potters’ marks in the Early Harappan period (Parpola 1996: 165).

During the Mature Harappan period the Indus Valley people continued to utilise the script and became highly skilled craftsmen, engineers and organisers. They built several vast urban centres, such as Harappa in the Punjab and Mohenjo-daro, which lies in Sind. These impressively large cities, with more than 40’000 inhabitants each, were

raised above the potentially detrimental flood level on massive artificial platforms of mud brick and show an astonishing level of planning and organisation with streets systematically laid out in a grid pattern (Rogers 2005: 201). Particularly the water-engineering projects of these sites, which include great watertight baths, hundreds of structurally robust wells and complex covered sewage-draining systems, are of a remarkably high standard unparalleled in that era (Parpola 2005: 30). The Indus people were not only adept at manufacturing bricks that would hold up constructions for thousands of years, they were also skilled manufacturers of ceramics and excellent metallurgists (Possehl 1996: 11). In addition, they produced intricately fashioned animal figurines and beautiful bead jewellery (Kenoyer 2006: 196-231).

These objects of the Indus Culture can be found throughout the Greater Indus Region. This territory includes southern Afghanistan, all of Pakistan, bar the most northern mountains, and it also extends into the northwest of India. All told, the Harappan Civilization covered an area of approximately a million square kilometres and the Indus people founded over a thousand settlements in this region. The Indus Civilization consequently qualifies as the largest civilization of its time, outranking in size the ancient empires of Mesopotamia, Akkad, Sumer and Elam (Possehl 1999: 155).

After about 2’000 BC, a general trend of deurbinisation set in as larger cities were being abandoned in favour of more modest agricultural villages (Weber 1991: 9-10). There was less social interaction between the different areas of the Greater Indus Region and the political and economic structure, which had prevailed throughout this territory, split up into smaller social and political entities (Kenoyer 2006: 26). As a result, areas of this sprawling civilization started to develop their own local material culture and traditions (Weber 1991: 9-10). Consequently, the Indus Culture began to disappear and was gradually replaced by new cultures, which were spreading into the region (Kenoyer 2006: 26). As the culture vanished, its script vanished with it. By 1’700 BC it ceased to be employed entirely and as time progressed the script and the mighty civilization, which had flourished for about 600 years, were eventually forgotten (Parpola 1996: 165).

Nowhere was it even recorded what language or languages the Indus people spoke. Sumerian, Egyptian, Elamite, Munda, Indo-Aryan and Dravidian are but a few of the languages, which have been proposed as the language underlying the Indus script

(Farmer / Sproat / Witzel 2004: 19). The Dravidian hypothesis would appear to have gained the greatest number of backers because the contemporary areal distribution of the Dravidian languages suggests that this language family might have stretched into the Indus Valley when the Indus Civilization was prospering (Salomon 1996: 746).

This is not the only popular suggestion however. A considerable amount of scholars have asserted that speakers of Indo-Aryan languages inhabited the Indian subcontinent earlier than has commonly been assumed. This claim has led them to the conclusion that the Indus inscriptions could possibly record an early form of Indo-Aryan (Farmer / Sproat / Witzel 2004: 21).

It also seems plausible that the Indus language family died out altogether and left no traces, which would allow linguists to reconstruct it. This last suggestion would make a decipherment of the Indus script nigh on impossible as an unknown script, which encodes an unknown language, cannot be successfully decoded (Robinson 2007: 145-148). Therefore, it is hoped that the Indus language was not entirely replaced and forgotten, like the civilization itself was forgotten.

1.2. Inscribed artefacts

The Indus Civilization remained a forgotten society until the 1920's, when the cities of Harappa and Mohenjo-daro were partially excavated. Only as these sites were being investigated, did anyone realize that a great civilization had occupied the Indian subcontinent before the Aryans had done so in the second millennium BC (Parpola 2000: 4-6).

Although the extent of the Indus Valley society was only grasped in the 20th century, an object bearing its script had already been unearthed and published in 1875. As the excavations of the major Indus cities progressed, a growing number of inscribed artefacts came to light (Joshi and Parpola 1987: XI). To date, over 4'000 of these objects have been recovered. Over half of them are small engraved stamp seals, the majority of which display a pictorial motif in a central position, accompanied by a line of symbols running overhead (Rogers 2005: 201). Wild, domestic and mythological animals, such as buffalo, bulls or unicorns are common seal images but more complex scenes, possibly portraying religious or mythological events, have also been found (Kenoyer 2006: 73).

These artistically rendered motifs could grant some insights into the type of religious customs practiced by the Indus people, as well as possibly aid with the decipherment of the symbols, for the inscriptions might refer to the pictorial content. Other indications, for the type of subject matter that might have been encoded on these artefacts, could be gleaned from their usage. Some of the intaglio-engraved steatite seals were used to make impressions on wet clay pots and clay tablets. These tablets may have been attached to shipments of goods (Shah and Parpola 1991: XXI). The inscriptions would thus serve as certificates of integrity and act as trade labelling devices, possibly specifying the contents, the dispatcher or the recipient of the merchandise (Kenoyer 2006: 88-89).

Apart from these pictorial motif seals an array of other artefacts bear the Indus script. There are, for instance, a number of rectangular tablets lacking any image and showing only a row of script symbols. Some of these tablets were fashioned in bas-relief, while others have been incised. These items may have served as votive offerings or perhaps as tokens of exchange used in commercial transactions (Wells 1998: 31-50). The system of symbols may also have been employed in a protective or otherwise magical manner, as suggested by the presence of the script on every-day objects, such as stone bangles, beads or tools (Kenoyer 2006: 76). These last objects are among the less frequently found inscribed artefacts, unlike ceramic vessels, which carry symbols more often. Pots were not only stamped with seals, they were also incised with Indus symbols and, on rare occasions, signs were even painted on pottery (Wells 1998: 33-34).

Entirely absent amongst these finds, are inscriptions on perishable materials, such as palm leaf, cloth or bark. Only durable substances such as stone, metal, terracotta, ivory and shell have survived the ages, which does not seem completely surprising, considering that the artefacts needed to endure for close to 4’000 years. However, the collection of more hardy inscribed objects does not include any monument-like items either, such as votive boulders or sculptures. The only inscription of monumental size is a three-metre wide, short string of symbols laid out on a floor in the city of Dholavira (Farmer / Sproat / Witzel 2004: 22).

1.3. Characteristics of the script

All these types of inscribed objects share the characteristic of carrying only extremely short messages. The shortest inscription bears only a single symbol, while the longest has a total of 28 signs spread across three sides of prismatic object. The longest text on a single side contains 17 symbols and the longest unbroken chain of signs is 14. These longer messages, which do not demonstrate any form of word division, are rather rare however, so the average Indus text contains a mere five signs (Parpola 1996: 166).

Researchers have generally agreed that the direction of writing is right to left for the vast majority of these texts. Overlapping strokes on clay and the spacing of symbols at seal peripheries support this claim. Where seals are concerned it must be taken into account that the right-to-left direction applies to their impressions (Witzel and Farmer 2000: 12). While the point of direction of writing has gained relatively wide acceptance, opinions divide rather sharply on a number of other aspects, such as the total amount of distinct signs and consequently also the type of writing system.

Estimates as low as 50 or 60 basic symbols have been proposed, which leads some scholars to advocate that the Indus writing system was of the syllabic or alphabetical type, i.e. a phonemic writing system (Rao 1982: 20-184). However, other estimates lie as high as 500 or 600 signs, which would exclude it as purely phonemic script (Wells 1998: 60-70). While extreme proposals such as these exist, a number of academics agree that 400 signs seems a more reasonable approximation (Possehl 1996: 54). If this amount is considered correct, the symbols are most likely part of a logo-syllabic writing system because a purely logographic script usually requires a greater number of symbols. The Indus signs therefore represent syllables on some occasions and whole words or concepts on others, which would explain how some inscriptions could be so brief. This hypothesis also ties in with the historical context, for contemporary scripts of Mesopotamia or Egypt also encoded language in this manner and the Indus people, who traded in those areas, might possibly have come across this script type (Parpola 2000: 84-85).

This is not to imply, that the Indus symbols themselves were adopted from any literate society, for the signs bear no resemblance to any known script, other than can be expected from symbols, which are iconic depictions of objects or basic geometric shapes. The Indus script contains simplified anthropomorphic and zoomorphic signs, in

addition to geometric forms, such as circles, squares and triangles, as do many other scripts, so accidental similarities are bound to occur (Robinson 2007: 147). Ligatures and composite signs also seem to be present amongst the Indus corpus, making the identification of distinct signs, i.e. graphemes, even more difficult.

The different graphemes are neither equally common nor do they occur in all text positions with the same frequency. Approximately a quarter of the symbols appear only once in the entire corpus, while 30 signs or so turn up between 100 and 500 times, and one symbol can be found over 1'000 times (Parpola 2000: 78-79). This sign, a 'U' with two horns protruding from either side, constitutes ten percent of all script data and, as a rule, it appears at the end of inscriptions rather than at the beginning. The same positional distribution applies to some other symbols as well (Parpola et al. 1969: 64). When sign sequences are analysed statistically, it also transpires that certain combinations of signs recur with a higher frequency than others and appear in a particular order too, making it seem like the inscriptions follow certain rules (Parpola 2005: 41-42).

Some researchers have taken this degree of order within the sign system, which is referred to as conditional entropy, as proof that the inscriptions show a linguistic structure (Rao et al. 2009: 1165). However, neither this phenomenon nor any of the other mentioned statistics and characteristics have been able to convince all researchers that the system of symbols at hand encodes a form of speech at all (Farmer / Sproat / Witzel 2004: 19).

1.4. Possibility of a non-linguistic symbol system

Some academics argue that the Indus signs might instead represent a non-linguistic symbol system, akin to heraldic blazons for instance. This challenging claim rests on several observations. Firstly, they note that conditional entropy also appears in non-linguistic systems, which leads them to refute the contention that the degree of order can be taken as evidence for linguistic structure (Farmer / Sproat / Witzel 2009: 1).

Secondly, they remark that the number of Indus signs and the statistics of their frequencies, which have been utilized as indicators for script type identification, are also compatible with statistics of non-script systems.

Thirdly, some researchers draw attention to the fact that repetitions of the same sign in a single inscription are far less common on Indus artefacts than on Egyptian cartouches. Considering that these two corpus samples show a comparable text length and supposedly match in script type, this seems an odd fact. This peculiarity leads them to infer that phonetic encoding could not have been as copious in the Indus script or might indeed be completely absent. This assertion would subsequently rule out the system being partly syllabic.

A fourth observation concerns the number of signs, which occur only once in the entire corpus. According to some scholars, this amount reaches levels that are unusually high and uncharacteristic for scripts, but are quite common in symbol systems that are not tied to language, such as the Vinca inscriptions. The Vinca inscriptions, which can be found on relatively crude southeastern European artefacts dating from the seventh, six and fifth millennia BC, were created by the cultures inhabiting Europe at that time and were presumably not a form of writing that encoded speech. Like the Indus script, the corpus of Vinca signs also spread over a large geographical area and contains standardised high frequency symbols as well as a vast amount of singletons (Farmer / Sproat / Witzel 2004: 22-39).

Next, scholars point out that no literate civilization that produced a significant body of inscribed objects ever failed to leave behind traces of longer texts. Therefore the unparalleled brevity of the Indus inscriptions raises the suspicion that long texts might never have been composed because their symbols did not encode language (Farmer / Sproat / Witzel 2004: 22-23).

The culmination of all these valid points, certainly make it seem possible that the Indus signs are not a form of writing, but a non-linguistic collection of symbols instead. However, none of the arguments prove conclusively that this is in fact the case (Parpola 2008: 111-119. Therefore, neither camp is presently in possession of any irrefutable evidence capable of settling the matter. Some of the main adversaries of the script theory did specify though, under which circumstances they would be inclined to change their position. They would consider the system a script if “remnants were discovered of an Indus inscription on any medium, even if imperfectly preserved, that contained clear evidence that the original contained several hundred sings” (Farmer / Sproat / Witzel 2004: 48). One single long Indus text would, therefore, elevate the Indus symbols to the

status of a script. A slightly shorter Indus inscription of approximately 50 signs would achieve the same effect, if some of the symbols on it were repeated in a manner typical of scripts (Farmer / Sproat / Witzel 2004: 48). Hence the discovery of a lengthy Indus inscription would be of great significance and value to the study of the Indus symbols.

2. The manuscript

An artefact with the potential to fulfil the second length requirement, and possibly also the first, has recently been located in an Afghani museum. A superficial glance at the symbols on this artefact reveals that some of the markings strongly resemble Indus signs. The first line of the text, for instance, contains symbols very similar to some of the most frequent and well-known Indus signs, such as the ‘horned U’ and the various fish-shaped symbols. Even with a very rudimentary knowledge of the Indus symbols, these signs catch ones eye immediately and could give rise to the assumption that the remaining less prominent characters might also bear resemblances to Indus signs when investigated more thoroughly.

A quick and unmethodical scan through the various lines of symbols lead to a rough symbol count of about 200 signs. This number of signs would definitely satisfy the minimum length condition, but it would remain to be seen if the inscription demonstrates the required sign repetitions. Meanwhile, a closer look at the edges of this artefact suggests that the object may consist of more than one inscribed layer. If this suspicion were correct, then the text length could also reach several hundred symbols, which would consequently also satisfy the first requirement of the Indus script opponents.

However, these perfunctory investigations need to be carried out in a more systematic and reliable manner. A careful analysis of the symbols on the artefact will therefore be necessary. Before this analysis is performed though, a description of the artefact itself will be given, as well as some general information pertaining to it.

2.1. General information

The artefact, which consists of a rectangular paper-like substance, with seven lines of symbols running from side to side, resides in the Sultani Museum in Kabul. This private

museum was founded only five years ago, by an Afghani gold trader by the name of Ahmad Shah Sultani, who managed to amass a considerable number of Afghani antiquities pertaining to various stages of history. The collection, which houses an estimated 3’000 objects, has not yet been properly catalogued, so some items lack precise descriptions and have, most probably, not been meticulously researched (Clammer 2007: 92).

Amongst these rather poorly documented artefacts is the inscribed object suspected of having Indus script connections. The label in front of the object, which reads: “An unspecific letter which has seven lines from skin bark”, supports the claim that little is known about the artefact (Western Himalaya Archive Vienna 2009: media identity number SN05 1000, 7085). No information on its age or the location of its discovery is known and no speculations as to the type of script have been put forward. The label merely states the number of text lines and concludes that the material in question is some type of bark.

The only possible indication, as to the date of the object, stems from the wooden box in which the inscribed object itself is resting. The lid of the box, which is propped up behind the artefact, has been decoratively painted with a scene depicting three riders on horseback, which appear to be brandishing clubs. The presence of horses would immediately lead to the conclusion that the box was not fashioned in the Indus era, as this animal was supposedly absent on the Indian subcontinent until the end of the second millennium BC (Parpola 2004: 157). The general style of the painting, the composition, the attire of the horsemen, the clubs, the flying ball and the two posts in the bottom left corner, are reminiscent of 16th century depictions of Persian polo games though. This could possibly indicate that the inscribed object dates from the same time period, which would not necessarily imply, that the text itself was not an Indus script composition. It would merely mean that the text was perhaps repeatedly copied over time from an Indus original. A second possible scenario involves the text being written with Indus symbols but in a non-Indus language. For some other people, who inhabited the Greater Indus region after the Indus civilization had fallen, may have adopted the form of the signs they saw on Indus relics. Another distinct possibility is, of course, that the box and the inscribed item are wholly unconnected. As no provenance is available the relationship between the two items can only be guessed.

A radiocarbon dating test would of course be able to ascertain the age of the artefact, but none could be performed in conjunction with this bachelor thesis. Therefore, it is also impossible to establish whether or not the object is in fact a forgery. For this reason the artefact will simply be assessed as if it were authentic and all scenarios remain open.

A carbon dating test was out of the question in this particular instance because no physical access was attainable, due to the remote location of the country in which the item is currently kept. This means that the entire thesis examination will be performed on the basis of three photographs, which were taken in the museum and are currently posted on the website of the Western Himalaya Archive Vienna (Western Himalaya Archive Vienna 2009).

The first picture shows the artefact, the wooden box and the abovementioned label (Figure 1), while the second picture is a close-up of the inscribed object, taken from a perpendicular angle to the supposed bark plane (Figure 2). This angle affords an excellent view of the inscription as practically all the symbols are in focus and discernable. This shot therefore represents the main source for the sign analysis, while the third photograph was only rarely used, due to the fact that large parts are blurred (Figure 3). As a result, this third shot only aided with the examination of certain areas of the lower object periphery.

2.2. Writing materials

The label visible on the first of these three photographs declares that the artefact standing behind it has been fashioned from bark, but an inspection of the material will reveal whether this assumption can be corroborated here.

The material consists of a thin beige coloured substance, which appears to be peeling apart into sheer layers, as can be observed particularly clearly in the bottom left corner of the artefact. Figure 4, an enlargement of the corner, shows this peeling effect, and simultaneously provides evidence that at least one of these layers was once completely separated from the uppermost inscribed layer, a fact that can be deduced from a lower level folio having been inscribed as well. An additional substance characteristic also revealed in that particular area, is the presence of a white film. The material therefore consists of a white layer as well as a light-brown one.

The entire beige plane, whose dimensions seem to come close to an A4 piece of paper, is covered in long red-brown ovals. These lens shaped marks, which are scattered in a parallel fashion, are only several millimetres wide and vary in length from approximately one centimetre to five centimetres. These types of spots only appear in the stems of woody plants, where they serve the purpose of diffusing essential gasses from the outside atmosphere to the core of the organism. So these spots, referred to as lenticels, mean that the artefact must indeed be of plant origin. The question remains which kind of plant (Sinclair 2000: 87).

The length and shading of these lenticels, their distribution pattern and the two-coloured layers with their peeling quality coincide with descriptions and pictures of birch bark, paper birch or silver bark in particular (Suryawanshi 2000: 2-5). This would therefore seem a justifiable suggestion, especially as birch bark has been utilised as a material for manuscripts in a number of regions, including northern India and Kashmir (Suryawanshi 2006: 103). Birch bark had been commonly employed in those areas during ancient and medieval times, until it was phased out and replaced in the 10th century by the introduction of paper (Suryawanshi 2000: 1). The oldest surviving remnants of a birch bark manuscript date from the second or third century, which would suggest that the inscribed artefact in Kabul could not very well be of any greater antiquity than these oldest examples, which were written in the Kharoṣṭhī script (Batton 2001: 21).

These bark items tend not to achieve any great age because they are susceptible to the effects of heat and moisture, which damage their tensile strength and flexibility, although in a far lesser degree than paper. Birch bark is one of the more resilient perishable objects, but the unforgiving monsoon climate of the Indian subcontinent eventually causes brittleness and succeeds in dissolving the natural adhesive holding layers together (Suryawanshi 2000: 5-7). This process of deterioration has begun in the Kabul manuscript, for not only are layers peeling apart, but small sections have crumbled away on the right-hand side and a long tear almost divides the bark into two pieces. It may be that an attempt at dividing two inscribed sheets lead to this damage, for folios often stick together and are troublesome to separate unless properly treated (Batton 2001: 22). This delicate state could explain why the second level of writing has not been presented in the museum. Fear of losing more of the text may also have

prevented a curator from removing dirt, dried mud and what appears to be vestiges of a root system or traces of some kind of fibers.

Despite this deterioration and covering residue, the black symbols remain astonishingly vivid, which would be astonishing, if a carbon-based ink, such as India ink, had been chosen. This type of writing fluid shows considerable stability and is faded little by sunlight (Nickell 1990: 35). Whether the ink was applied with a pen or a brush seems a little unclear though.

The symbols certainly do not exhibit the tell tale signs of split-reed pen texts, for at no point does an inkless line appear in the centre of any stroke. This line would normally be created by the cut in the pen nib and would appear as the implement was running low on ink, which happens frequently with pens lacking in-built reservoirs (Salomon 1999: 108-109). However, a pen with an internal ink well would not require refilling as often as a dip pen and would not necessarily show this phenomenon, which is known as ink failure (Nickell 1990: 18).

The sequences of signs on the Kabul manuscript were carefully examined for any signs of ink failure and while some characters certainly fluctuate in thickness and intensity of colour, there do not seem to be any instances of progressively waning thickness followed abruptly by a bolder stroke. There are clear instances where the colour intensity weakens progressively as, for instance, on the left-hand side of the fist line, but unfortunately this could be a deceptive visual effect, for not only does the ink fade in this spot, but the colour of the bark itself fades too. So, the entire area might have been effected by sunlight, water or worn away by an abrasive force. Also, the dust and dirt, which rests on the artefact, can brighten the tone of the ink in places, which would cause signs to appear less vibrant. Therefore, colour variations on the bark are unreliable ink failure indicators.

The inability to locate any obvious ink failure could mean that the Kabul manuscript was written with a reservoir pen. This type of pen purportedly existed as early as the 10th century in the Islamic world (Nickell 1990: 18). However, it also seems possible that a brush was used because this implement does not require re-dipping as often as some dip pens.

A closer look at the individual strokes does not aid in settling the matter unfortunately, for some signs have thin lines running in one direction, while a thicker

stroke has been placed at a right angle to this line, which occurs when a pen is held at a particular angle. This pattern does not repeat completely systematically though, which could either indicate that the penman occasionally rotated the pen, or it could also mean that a brush was employed, for a similar effect can be achieved with this implement (Salomon 1999: 108-109). The question of the type of writing implement can therefore not be satisfactorily answered here, but fortunately this has little ramifications for the grapheme analysis. Establishing the direction of writing on the other hand is a preliminary step that is of consequence when trying to decipher a script.

2.3. Direction of writing

It is likely that the Kabul manuscript should be placed as it has been done in the museum, with the rows of signs running horizontally and the symbol consisting of three single strokes standing in the upper left-hand corner. This orientation is considered appropriate because some of the signs resemble anthropomorphic figures and it seems natural to place these stick-men in an upright position, not on their sides or on their heads. Also, the few immediately recognisable Indus-like symbols would be standing in their usual scrip orientation if this placement were adopted. The signs could have been rotated of course, but this seems less probable. With the orientation fairly firmly established, the direction of writing remains to be determined.

The arrangement of the symbols into horizontal strips, with spaces of only a few millimetres to either side, but a gap of approximately one and a half centimetres above and below, lead to the conclusion that these lines are meant to be read either from right to left or vice-versa. Reading from top to bottom would not be possible as the signs do not line up vertically in way that would make it obvious which symbol was to follow the previous.

While these points are fairly easy to establish, ascertaining the precise direction of writing becomes more challenging and problematic. Several aspects can be investigated, such as the colour intensity and thickness of the symbols for instance. Bold strokes followed to the left by increasingly thinner, paler and more ragged lines that abruptly return to bold strokes, would, for instance, imply a right-to-left direction of writing. But, as was mentioned earlier, the intensity and thickness of strokes reveal little in the case

of the Kabul manuscript. As this criteria failed to help, symbol spacing will be taken into account next.

The margins of inscribed objects are often inspected because crowding of signs at one particular periphery indicates that writing commenced at the opposite end and that the penman ran out of space as he approached the margin. Unfortunately the right edge of the Kabul manuscript is partially obscured by a strip of torn bark and has also been badly damaged so it cannot be used to investigate cramming. However, the left margin of the third, fourth and seventh line appear to show signs of this kind of symbol cramping, particularly the last two or three symbols of these lines.

The third line contains a wheel-like graph, which turns up repeatedly throughout the rest of the manuscript. In the other areas the wheel symbol takes on a fairly roundish shape, while in the third line it has become very elongated, suggesting that not enough remaining space was available to give it its usual form. The fourth row contains a sign with a similar story. This sign, a rectangle with a horizontal bar through its centre, also appears frequently in the text in a wider version, but has been reduced in size at the left periphery. This reduction is also observable in the last line, where three or four characters appear to have been shrunken down to an uncharacteristic height. It should be noted, however, that the bark is curling upwards in this section, so a trick of perspective may be partly responsible for this effect. While this last example may therefore be open to debate, the other instances seem to demonstrate that size reduction was necessary in order to fit symbols onto a particular line.

By contrast, there also appear to be some instances where the signs on the left take up more space than elsewhere. The left most symbol of the sixth row, for example, has been drawn a little wider than graphs of a similar general shape. The same applies to the last complete symbol in the second line. This horizontal spreading may indicate that the writer was attempting to fill the area so as not to leave a ragged margin when drawing the last sign in a row. This spreading and cramming at the left periphery would therefore suggest that the direction of writing was right-to-left, but these few observations alone cannot be taken as irrefutable evidence. For this reason any further text analysis would have to operate under the assumption that the opposite direction of writing, or indeed the boustrophedon style, were still distinct possibilities.

2.4. Graphemic Analysis

This investigation of the writing direction merely represents a preliminary step in the analysis of a text. The main part involves a graphemic analysis, which aims at establishing the number of distinct signs, which are referred to as graphemes. This in turn entails being able to distinguish graphemes from variations of the same sign. Before this step can be attempted though, it must first be determined how the markings on the manuscript are to be divided into individual graphs (Parpola 2004: 68). A segmentation of the text will therefore be attempted next.

2.4.1. Segmentation

The manner of segmentation seems quite clear in the majority of cases because most symbols consist of contained units, such as circles or triangles. These units have obvious boundaries and are separated from their neighbouring graphs by appreciable gaps. To illustrate which divisions have been deemed reasonably apparent, a hand-traced reproduction of the text has been created and partitioned into symbols (Figure 5). Black areas of this reproduction represent ink marks on the first layer of bark and red patches of colour denote ink lying on a lower sheet, while the blue line outlines the border of the top sheet. The fine vertical lines between the black areas show the actual segmentation. The resulting wedges have been numbered from right to left. The sign standing on the farthest right in the top line therefore receives the code 1.1, denoting that it is to be found in the first line, in the first position, while the first sign in the second line from the top is labelled 2.1 and so on. This allows an identifying digit to be assigned to every proposed graph.

These proposals should not be regarded as indisputable, for there are occasions where a definite division seems problematic. Several factors can be blamed for these instances of doubt. Firstly, dust, dirt and fragments of birch bark cover certain areas, thus concealing the signs beneath. Secondly, portions of bark have disintegrated in places, principally the edges and the lenticels. Lenticels are more prone to deterioration due to their chemical composition (Suryawanshi 2000: 4-7). These damaged or covered zones are portrayed by grey shading on the reproduction. Where the obstruction is so severe that barely anything can be made out, as in area 4.3-4.5, the fine segmentation lines have been dispensed with. In their stead short vertical bars have been placed,

indicating that the obscured area probably contains X amount of segmentation lines. The amount can be deduced from the size of the obscured area and the average amount of signs, which would generally take up that space. Using this method 36 signs are estimated to be obscured but the actual number could lie as much as a dozen or so higher or lower. The identification number of these 36 graphs is written in blue print so they can be recognised more easily in Table 5.

Segmentation also becomes complicated when signs appear that do not constitute contained units. Area 1.31-1.32, for instance, includes three individual strokes, which are separated from each other by small gaps. It therefore becomes difficult to gage whether these represent three individual symbols or two or maybe even a one single symbol. Occasionally hints can be obtained from the rest of the text. For example, if the same combination of strokes turns up many times, but the simple strokes never or rarely appear alone, it can be assumed that they should be analysed as one graph. One might hesitate analysing the three strokes of area 1.19 as one sign, for instance, but as this constellation of marks recurs at least four more times, it can safely be assumed that they do in fact form one graph.

Unfortunately area 1.31-1.32 is not quite as clear-cut, for the middle zigzag stroke only occurs with a diagonal stroke to its right, but the diagonal does not always face the same way. The diagonals do not show up alone in any other position. This could mean that the two diagonals were separate signs that only rarely occur, or it could mean that the penman did not need to care which way the diagonal was facing because the meaning would still have been obvious with a leftward or a rightward slanting diagonal as it was part of the zigzag sign. Either option seems possible, but the slightly more plausible was chosen for Figure 5.

The tentativeness of the proffered segmentation is signalled in Figure 5 by using red identification numbers instead of black ones. Their possible alternative interpretations have then been arranged in Figure 6.

There are 27 cases of this type of relatively uncertain graph proposal, which together with the 36 assumed obscured symbols and the 167 relatively clear cases, adds up to a total of 230 graphs. If all extreme cases of doubt are calculated together, the minimum number of graphs sinks to about 211 and if the maximum amount is tallied, there are approximately 267 graphs.

As can be gathered from Figure 5, there must also be additional symbols on a lower sheet of bark, for the ink marks in the uppermost section of the top line cannot represent continuations of the line 1 characters. They simply do not match up in a convincing manner. These partial symbols, which have been rendered in red on the reproduction, run almost the entire length of the upper artefact rim and several partial symbols can also be made out along the left rim of the manuscript. For instance, the area to the left of lines 3, 4, 5 and 7 shows evidence of writing, which does not pertain to the top inscribed layer. This would suggest that the entire lower sheet was filled in the same way as the top layer, implying that the total number of graphs could reach doubled the initial count of approximately 230 graphs. Whether there might even be a third or a fourth inscribed sheet remains unclear. The segmentation procedure must therefore be constricted to the top level.

2.4.2. Grapheme and allograph identification

The graphs having been established by the process of segmentation, it must now be ascertained how many different graphemes occur in the sample. This involves comparing the shapes of the signs and judging whether the differences between them are big enough to warrant the graphs being classified as distinct signs. If the forms are identical or very similar to each other, then the graphs are categorised as allographs, that is to say, variations of the same symbol (Parpola 2004: 68-89).

To facilitate the comparison, it can be beneficial to arrange the graphs in such a manner that all similar shapes appear close to each other. This organization has been attempted in Figure 7. The linear nature of the layout does not always allow all similar graphs to be positioned in the vicinity of each other, but the lack of other viable possibilities makes this the best available option.

The compilation of signs begins with the simplest forms, such as single vertical strokes, and continues with three-sided geometrical shapes and rectangular objects. The more complicated shapes, resembling humans and animals, are placed near the end, along with a few other miscellaneous shapes. In total, this collection lists 172 graphs, which means that 58 signs of the 230 segmented elements were deemed too damaged or obscured to be included in any further graphemic analysis.

In a next step, the list was divided into groups of allographs and every one of these groups was then allocated a number in ascending order. Graphs to the right of an allotted number are therefore considered variations of the same grapheme and the graph located to the immediate right of such a number is regarded as the most typical version of the various allograph shapes. As can be gathered from the last such number in Figure 7, 62 distinct graphemes have been posited for the Kabul manuscript.

Once again though, not all of these graphemes have been submitted with the same amount of conviction. Some situations require a considerable amount of conjecture and could well be analysed in different manner. For instance, the single stroke graphs, which have been separated into sign number 1 and 2, could alternatively be categorised as three separate graphemes, a very short apostrophe-like sign, a long vertical line and a medium length stroke. A continuum of progressively longer single strokes appears in the data and it seems unclear whether the intermediate length was intentionally created or whether it occurred because of an imprecise hand movement.

Incidents of inaccuracy would, after all, be expected in a hand-written document because controlling a writing implement and keeping the hand steady requires a certain amount of effort and concentration, which eventually dwindle. It is hard to speculate however, which minor variations are due to a lapse in concentration and which are meant to set graphemes apart. Slight differences in stroke length could result from either possibility, as could small deviations in stroke positioning.

Group 13, for example, contains arrow shaped graphs, which all have a similar curving line jutting out to the right. In some cases, the line cuts through the arrows triangle, while in others the line only just touches one side of the triangle. In one instance, the curve is even completely detached from the arrow. This varying placement could have resulted from human imprecision or been deliberately created to distinguish graphemes from each other. The same observation also applies to the groups 41 and 44.

An imprecise or hasty penman might also cut corners or simplify some of the more complex symbols. The swift execution of a rectangular shape, for instance, often leads a draftsman to produce blunt angles, or even rounded edges, instead of pointed angles because it takes less effort to produce the former. This could explain why a rounded variant of the otherwise diamond-shaped sign number 26 appears in the data.

A further simplification, often employed during the development of a writing system, involves the reduction of the number of strokes. Complex signs, composed of a great number of lines, are often reduced to fewer strokes to save time when writing them (Parpola 2000: 38). This could account for the modifications in group 23. This sign has been drawn with either two or three horizontal bars inside a rectangular box, which adds up to a total of six or seven strokes respectively. The execution of this fairly high stroke sum would consume a significant amount of time and would merit simplification. One horizontal bar might therefore have been omitted, in order to proceed a little more quickly to the next symbol. That the sign was only occasionally simplified may be explained by space considerations, for the writer sometimes seems to have drawn the box a little too small to allow three horizontal bars to fit into it and was forced to reduce the amount to two bars. Of course, this difference in stroke number could also be the distinguishing feature between two graphemes, as could the angular versus round variation.

To signal that these and a few other allograph categorisations are only of a tentative nature, red print has been employed in Figure 7. Two red graph identification codes standing next to each other indicates that a grapheme division could potentially be posited between the two graphs to which the codes pertain. If all these divisions were implemented, then 14 more signs would join the 62 already proposed graphemes. Therefore, the maximum of Kabul signs would reach 76 graphemes.

3. A comparison

3.1. Grapheme comparison

Now that the graphemic analysis of the Kabul manuscript has been concluded and a sign list has been put forward, a comparison between the Kabul sign list and the symbols of the Indus script can be launched.

In order to proceed more swiftly, it was considered prudent to compare the birch bark signs with an Indus sign list, which had already been prepared by an expert, as opposed to attempting a graphemic analysis of the Indus corpus from scratch. Parpola’s sign list was settled upon as it offers several advantages (Parpola 2000: 70-78). Firstly, Parpola’s list has been arranged in a compact and easy to grasp manner. Secondly, the

compilation does not just show what he suggests are graphemes but also shows what he regards as principal variants of a single grapheme. This means a researcher is not merely presented with a stylised version of what Parpola, or any other author, regards as the most typical form of several different variations of a sign, but gets a complete, or at least a near complete, overview of shape deviations that occur in the corpus. This allows the researcher to consider the validity of the proposed categorization of signs into graphemes and allographs. Thirdly, a researcher can check whether the stylisation of a graph is a faithful rendering of the original form because Parpola has included the identification code of the Indus artefact on which the graph occurs. Fourthly, the reproduction quality of the signs is higher than in some other publications because the chosen graph size is larger and the line weight is lower, which allows for a more detailed depiction (cf. Fairservis 1992: 149-152, Hunter 1993: 203-210 and Krishna Rao 1982: Figure 22-31).

An inspection of this list reveals that 26 of the 62 Kabul graphemes show a near identical resemblance to graphs in Paropla’s inventory. These 26 Kabul graphs and their Indus correspondences have been juxtaposed in Figure 8. Minor deviations do occur in this collection but the differences are negligible as they are of a systematic nature. Kabul graphemes 14, 30, 43, 55, 56 and 58, for instance, all show more angular shapes than Parpola’s depictions, and graphemes 36 and 38 have been drawn in a more disjointed manner than their supposed Indus counterparts. However, some of these variations do occasionally appear in the corpus of Indus inscriptions and the differences in the Kabul manuscript could be attributed to a general style preference of the birch bark penman (Shah and Parpola 1991: M-759 a, M-762 a, M-1601 A).

One additional sign, which is not present in Parpola’s list, has been added to Figure 8 because an inspection of the Indus inscriptions and two additional Indus sign inventories showed that it does in fact occur in the Indus corpus (Shah and Parpola 1991: 372, Krishna Rao 1982: Figure 29 and Wells 1998: 168). This finless fish-shaped symbol occurs predominantly on Early Harappan pottery and was probably not included in Paropla’s data because he qualified it as a pre-script potter’s mark (Shah and Parpola 1991: 372 and Joshi and Parpola 1987: K-110 A and K-111 A). If this graph were regarded as a known Indus symbol however, then there would be 27 Kabul graphemes bearing an extremely strong resemblance to Indus signs.

This number could be expanded to 36 graphemes because nine Kabul graphemes are identical mirror images of signs in Parpola's list, as can be gathered from Figure 9. All Kabul graphemes in Figure 8 have symmetrical shapes and would appear unaltered if they had been mirrored along a vertical axis. Therefore, it would seem plausible that all Indus signs have been flipped vertically in the Kabul manuscript. The reason for this axial mirroring of Indus symbols can only be speculated upon, however, reversing the orientation of signs and the direction of writing was not a wholly unusual phenomenon in the Indus corpus, nor in other ancient scripts for that matter, as writing practices were not as rigidly set and followed in ancient times (Kenoyer 2006: 72). The manuscript penman may have simply adopted a variety of the Indus script, which preferred the non-canonical orientation of symbols.

Whatever the reason may be for the reversal of the canonical symbol orientation, the change could be considered systematic, as could the modification of the signs in Figure 10. These eight graphemes show strong similarities to Indus symbols but lack one stroke or more. This reduction of lines is a common form of writing simplification, as was mentioned earlier, and it would have been a necessary measure in some of these instances because the writing implement would not have been delicate enough to create all of the customary lines on such a restricted space. Signs 15 and 34 would not have room to fit in another horizontal bar for instance.

A further four Kabul graphemes are suspected of sharing a strong similarity to Indus signs but as they are partially obscured the correspondences are tentative. For this reason they have not been grouped with the graphs in Figure 8, but have been listed separately in a Figure 11.

Figure 12 contains four Kabul graphemes, which seem to bear a fair resemblance to Indus signs and Figure 13 shows seven signs, which could be regarded as slightly modified versions of already listed Kabul graphemes. For example, grapheme 31 is grapheme 32 without the central arrow and grapheme 59 is a combination of sign 55 and 58. These signs do not appear in the corpus in this exact form, but necessarily resemble Indus symbols because they are only mildly dissimilar from signs, which have been established as near identical Indus symbols. In addition, two of these examples exhibit modifications, which do turn up in the Indus corpus. Brackets occur around certain Indus symbols for instance, as they do around Kabul sign 57 and there is also a

stick figure with a horizontal line running through it like in Kabul grapheme 49 (Parpola 2000: 70-71).

The last grapheme in Figure 13, Kabul sign 12, has been set apart from the other examples because one Indus sign inventory lists a symbol with an identical shape, but it could neither be located in the corpus of inscriptions nor be found in any other available inventory (Krishna Rao 1982: Figure 24). As it's existence could not be verified by a second source, it was deemed too uncertain to be included in the group of identical correspondences.

If all these near identical and similar matches are added together it transpires that only three of the 62 Kabul graphemes do not appear to resemble any known Indus symbol. These three signs are given in Figure 14.

3.2. Comparison of alternative graphemes

So far only the most probable Kabul graphemes have been compared to Indus symbols, but it seemed sensible to check whether any of the discussed alternative segmentations or allograph categorisations would match any Indus signs.

Figure 15 shows the alternative segmentations, for which a similar Indus signs could be found. Some of these similarities do seem as plausible as their previously suggested options, but as none of the new possibilities appeared considerably more credible, the segmentation was not revised.

Figure 16 lists the allograph pairs in the Kabul data, which have matching forms in the Indus corpus. Parpola categorises three of these allograph pairs as variants of three different Indus graphemes. This would support the original Kabul grapheme division that was suggested in Figure 7. The two remaining pairs in Figure 16 have been classified as four separate graphemes by Parpola (Parpola 2000: 70-78). This would call into question the grapheme divisions in this thesis. However, two of the signs in this group do not have identical or near identical Indus correspondences. The matches are only similar, so these allograph classifications were not modified either.

Considering the alternative segmentation and allograph possibilities after a comparison with the Indus corpus, therefore, did not drastically change the previous view of the Kabul text analysis. It will still be assumed that there are approximately 62

Kabul graphemes and that 59 of these signs show similarities with symbols in the Indus corpus.

This degree of matching correspondences strongly implies that the two grapheme lists must be related to each other. It seems highly unlikely that a non-Indus society could have created such a large number of strikingly similar forms without having been acquainted with some Indus inscriptions. However, it has been noted that 40 to 50 Indus graphs also bear a strong resemblance to symbols present in the writing system of Easter Island and it is thought that the inventors of this system could not have been familiar with the Indus script. This shows that unconnected societies can occasionally produce similarities of this extent (Robinson 2007: 147).

3.3. Comparison of sign frequencies

A considerable amount of sign correspondences would therefore not conclusively prove that two writing systems were related. However, if the individual signs also occurred with a comparable frequency in the two scripts, then it could be convincingly ruled that the grapheme inventories had been thought up separately, for no unrelated scripts would share greatly similar sign frequencies. The Kabul manuscript signs and the Indus symbols could therefore be proven to share a connection if a large number of the most frequently occurring Indus graphemes also appeared on the sheet of birch bark.

In order to investigate this possibility, the top 60 most common Indus signs were searched for in the Kabul data. It was discovered that nine of the ten most common Indus graphemes and 17 of the top 20 Indus graphemes occur in the manuscript. When more signs are included in the comparison it becomes clear that the total continues to increase but less steeply. 23 of the top 30, 27 of the top 40, 30 of the top 50 and 35 of the top 60 Indus symbols are present on the Kabul birch bark (Parpola 2000: 70-78 and Wells 1998: 60-70). Figure 17 shows the top 60 Indus graphemes. Signs, which were not found in the manuscript, have been marked with red print.

In comparison only two of the top ten, four of the top 20, ten of the top 30, 11 of the top 40, 14 of the top 50 and 16 of the top 60 Indus signs appear in the Easter Island script (Possehl 1996: 94-95). This means that the Easter Island script has a considerably lower total of frequent Indus signs than the Kabul text and it also shows that the total

increases far less steeply in the collection of Easter Island symbols. These statistics have been graphically summarized in Figure 18.

These results would seem to imply that the Kabul signs are indeed related to the Indus script while the Easter Island script is not. The Easter Island total is too low while the total number of Indus symbols on the Kabul manuscript is still in the acceptable range. The Kabul total would not be expected to reach a considerably higher amount than 35 because only 172 graphs are visible on the bark fragment. Not all the 60 most common Indus signs would therefore be expected to appear in such a relatively short text.

This frequency argument would gain even more strength if the most common Kabul graphemes also occurred in the 60 most often repeated Indus signs. An exploration of this matter reveals that the most frequent Indus sign is also the most common Kabul grapheme and shows that ten of the top 11 most common Kabul graphemes appear in the list of the 60 most frequent Indus signs, as can be gleaned from Figure 20. When less common graphemes are added to the frequency comparison, the amount increases less steeply once again. 14 of the top 16, 18 of the top 25 and 23 of the top 37 Kabul signs are amongst the 60 most common Indus signs, as can be seen in Figure 19. This would also support the claim that the sign shapes of the two grapheme lists were related to each other.

4. Conclusion

Whether this last statistic also indicates that the signs encode similar phonetic and logographic values in the two systems, remains unclear however. It just seems too problematic to judge with any certainty, whether the Kabul text could exhibit this pattern of sign frequency overlap, if the symbols represented entirely different linguistic units in the Indus corpus. One would need to decide, whether sign frequencies could coincide in this manner, if a non-Indus people had stumbled upon a collection of Indus inscriptions and merely adopted the shapes of some of the most common symbols for their script without adopting their encoding value. If this scenario could be discounted on the basis of signs frequencies, then it could be assumed that a number of Kabul graphemes share the same value as their matching Indus symbols. This, in turn, would

imply that the Indus symbols belong to a speech-encoding writing system and the hypothesis of the non-linguistic Indus symbol system would be refuted.

If this problem were solved, then researchers would need to establish whether the Kabul text employed the exact same writing system as the Indus Civilization or whether the system was modified in any way. The manuscript might represent a later simplified stage of the Indus script for instance. This hypothetical stage could contain a reduced number of logographs and rely predominantly on syllables instead. Evolutionary patterns of other scripts show that this development is a widespread phenomenon (Parpola 2000: 38).

A large portion or maybe even all of the Kabul signs might therefore encode syllables. The amount of proposed Kabul graphemes would not rule out either of these possibilities because the total of Kabul graphemes might increase drastically, or merely a little, if a larger corpus than 172 graphs were available for examination. This relatively limited corpus simply does not allow predictions on the matter. Therefore, it would certainly be worthwhile uncovering and analysing the other layers of bark to find out how many more graphemes would join the 62 already visible graphemes.

It would also be extremely interesting and potentially revealing to compare the frequencies of particular signs sequences. If it could be demonstrated that the most common Indus grapheme sequences also occur relatively frequently in the manuscript, then it would seem likely that the Kabul penman not only used the Indus script but that the text encoded the language, or one of the languages, of the Indus Civilization. As it stands the manuscript might encode a non-Indus language and merely utilise the Indus writing system.

Many issues therefore still remain unresolved and numerous aspects beg further investigation. So far, a graphemic analysis of the Kabul manuscript has only enabled the positing of a Kabul grapheme list and its subsequent comparison to the symbols of the Indus corpus has merely lead to the definite conclusion that the majority of the signs on the analysed birch bark were borrowed from the collection of Indus symbols. It should also be kept in mind that any of the contentions that were presented in this thesis would be invalidated if the Kabul manuscript proved to be a forgery. As all arguments hinge on the assumption that the manuscript is not a fake, establishing the date of the strip of

birch bark would certainly head the list of top priorities, which need to be addressed in any further examination of this intriguing artefact.

References

Batton, Susan Sayre

2001 "Trennungsangst: Das Auseinandernehmen und Konservieren eines buddhistischen Birkenrinden-Manuskript.", in: *Restaurator*. Vol. 107. No 1: 20-24. Copenhagen: Munksgaard.

Clammer, Paul

2007 *Afghanistan*. London: Lonely Planet.

Fairservis, Walter A.

1992 *The Harappan Civilization and its Writing*. Leiden: E. J. Brill.

Farmer, Steve, Richard Sproat and Michael Witzel

2004 "The Collapse of the Indus Script Thesis: The Myth of a Literate Harappan Civilization." *Electronic Journal of Vedic Studies*, 11: 19-57 [URL: http://www.safarmer.com/fsw2.pdf, 27 April 2009].

Farmer, Steve, Richard Sproat and Michael Witzel

2009 "*A Refutation of the Claimed Refutation of the Nonlinguistic Nature of Indus Symbols: Invented Data Sets in the Statistical Paper of Rao et al.*" [URL: http://www.safarmer.com/Refutation3.pdf, 27 April 2009].

Hunter, G. R.

1993 *The script of Harappa and Mohenjodaro and its connection with other scripts*. New Delhi: Munshiram Manoharlal Publishers.

Jarrige, Jean-François

1988 "Introduction", in: *Les cites oubliées de l'Indus*. Paris: Association Française d'Action Artistique.

Joshi, Jagat Pati and Asko Parpola

1987 *Corpus of Indus Seals and Inscriptions, Vol. 1: Collections in India*. Helsinki: Suomalainen Tiedeakatemia.

Kenoyer, Jonathan Mark

2006 *Ancient Cities of the Indus Valley Civilization*. Oxford: Oxford University Press.

Krishna Rao, M. V. N.

1982 *Indus Script Deciphered*. Delhi: Agam Kala Prakasham.

Lawler, Andrew

2004 “The Indus Script – Write or Wrong?”. in: *Science.* Vol. 306. 17 December 2004: 2026-2029 [URL: http://www.sciencemag.org/cgi/content/summary/306/5704/2026, 13 November 2007]

Nickell, Joe

1990 *Pen, Ink, and Evidence: A Study of Writing and Writing Materials for the Penman, Collector, and Document Detective.* Lexington: The University Press of Kentucky.

Parpola, Asko

1996 “The Indus Script.”, in: *The world’s writing systems*. Edited by Peter T. Daniels and William Bright. New York: Oxford University Press. 164-171.

2000 *Deciphering the Indus Script*. Cambridge: Cambridge University Press.

2005 *Special Lecture: Study of the Indus Script*. Paper read at the 50th ICES Tokyo Session. 19 May 2005 [URL: http://www.harappa.com/script/indusscript.pdf, 27 April 2009]

2008 “Is the Indus Script Indeed not a Writing System?”, in: Airāvati. 111-131. Jawahar Nagar, Chennai: Varalaaru. [URL: http://www.harappa.com/script/indus-writing.pdf, 27 April 2009]

Parpola, Asko et al.

1969 *Decipherment of the Proto-Dravidian Inscriptions of the Indus Civilization.* Copenhagen: Scandinavian Institute of Asian Studies.

Possehl, Gregory L.

1996 *Indus Age: The Writing System*. Philadelphia: University of Pennsylvania Press.

1999 *Indus Age: The Beginnings*. Philadelphia: University of Pennsylvania Press.

Rao, Rajesh P. N. et al.

2009 “Entropic Evidence for Linguistic Structure in the Indus Script.”, in: *Science.* Vol. 324, 23 April 2009: 1165 [URL: http://www.cs.washington.edu/homes/rao/ScienceIndus.pdf, 20 July 2009]

Rao, Shikarpur Ranganath

1982 *The decipherment of the Indus Script*. Bombay: Asia Publishing House.

Robinson, Andrew

2007 *The Story of Writing*. London: Thames and Hudson.

Rogers, Henry

2005 *Writing Systems: A Linguistic Approach.* Oxford: Blackwell Publishing.

Salomon, Richard

1996 “Deciphering the Indus Script: Review Author: Richard Salomon.”, in: *Journal of the American Oriental Society*. Vol. 116. No. 4: 745-747.

1999 *Ancient Buddhist Scrolls from Gandhara*. Seattle: University of Washington Press.

Shah, Sayid Ghulam Mustafa and Asko Parpola

1991 *Corpus of the Indus Seals and Inscriptions, Vol. 2: Collections in Pakistan*. Helsinki: Suomalainen Tiedeakatemia.

Sinclair, John M.

2000 *Collins English Dictionary*. Fifth edition. Glasgow: Harper Collins Publishing.

Suryawanshi, D. G.

2000 “An Ancient Writing Material: Birch-Bark and its Need of Conservation.”, in: *Restaurator*. Vol. 21. No. 1: 1-8. Copenhagen: Munksgaard.

Suryawanshi, D. G.

2006 “Like Paper: Birch Bark and its Chemical Composition.”, in: *Restaurator*. Vol. 27. No. 2: 103-113. Copenhagen: Munksgaard.

Weber, Steven

1991 *Plants and Harappan Subsistence*. Boulder Colorado: Westview Press.

Wells, Bryan

1998 *An Introduction to Indus Writing*. MA thesis. Department of Archaeology. Alberta Canada: University of Calgary [URL:

http://dspace.ucalgary.ca/bitstream/1880/25900/1/31309Wells.pdf, 27 April 2009].

Witzel, Michael and Steve Farmer

2000 "Horseplay in Harappa: The Indus Valley Decipherment Hoax.", in: *Frontline*. Vol. 17. 13 October 2000: 4-14 [URL: http://www.safarmer.com/frontline/horseplay.pdf, 27 April 2009].

Western Himalaya Archive Vienne

2009 Susy Novotny media creator [URL: http://whav.aussereurop.univie.ac.at/ic/151/?page=4, 8 November 2009].

Appendix

Figure 1. Photograph of the artefact with wooden box and label. Media identity number SN05 1000, 7085 in: Western Himalaya Archive Vienna 2009.



Figure 2. Photograph of the artefact close-up. Media identity number SN05 1000, 7087 in: Western Himalaya Archive Vienna 2009.



Figure 3. Photograph focused on the lower portion of the artefact. Media identity number SN05 1000, 7086 in: Western Himalaya Archive Vienna 2009.



Figure 4. Partial enlargement of bottom left-hand corner of media identity number SN05 1000, 7086 in: Western Himalaya Archive Vienna 2009.

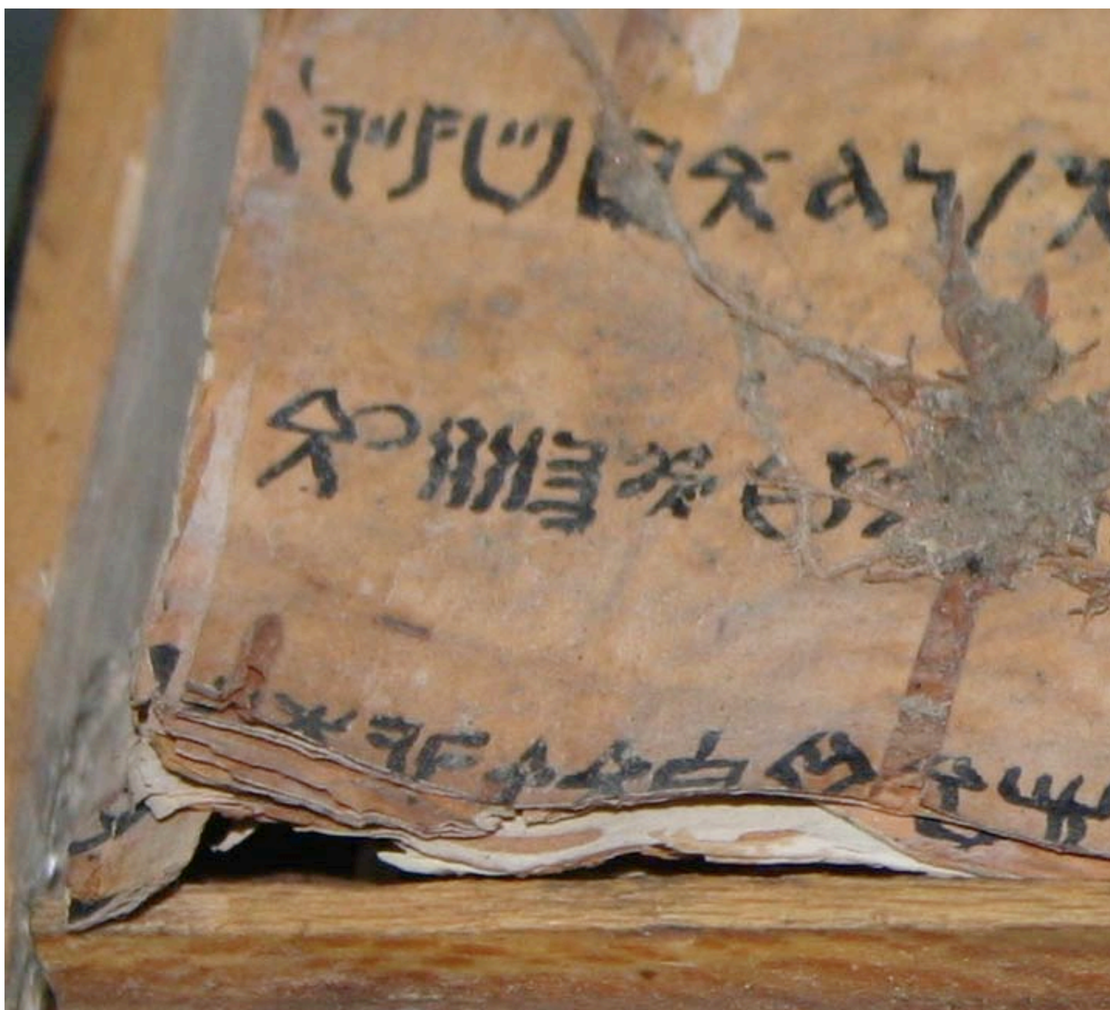
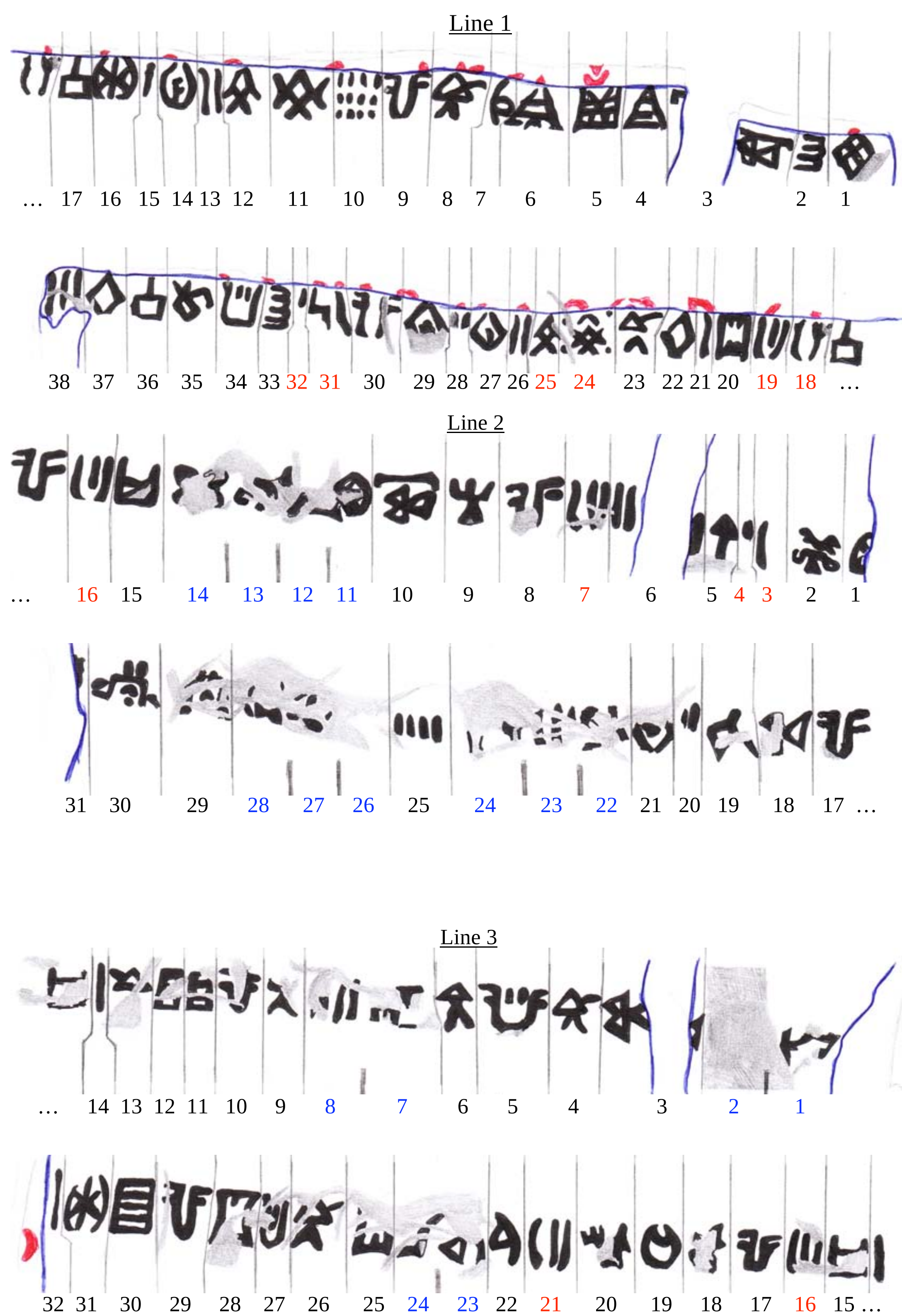
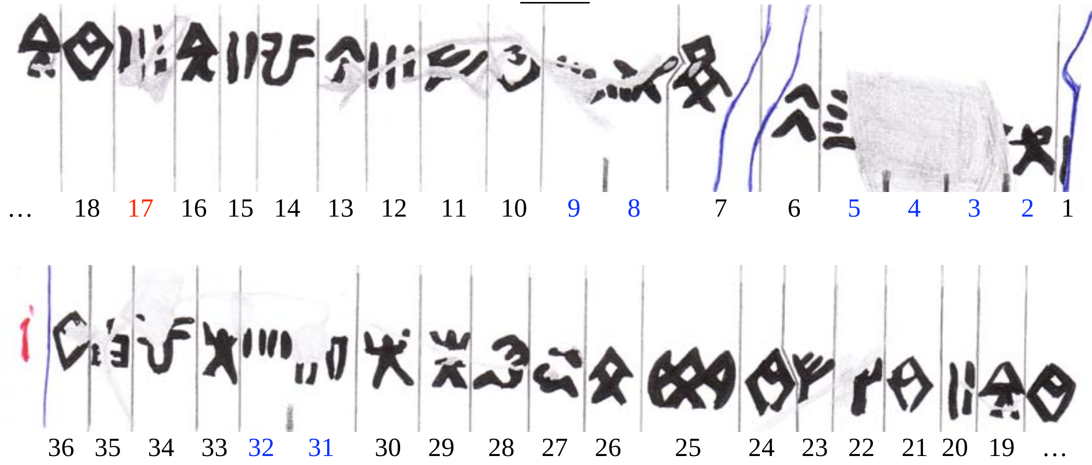


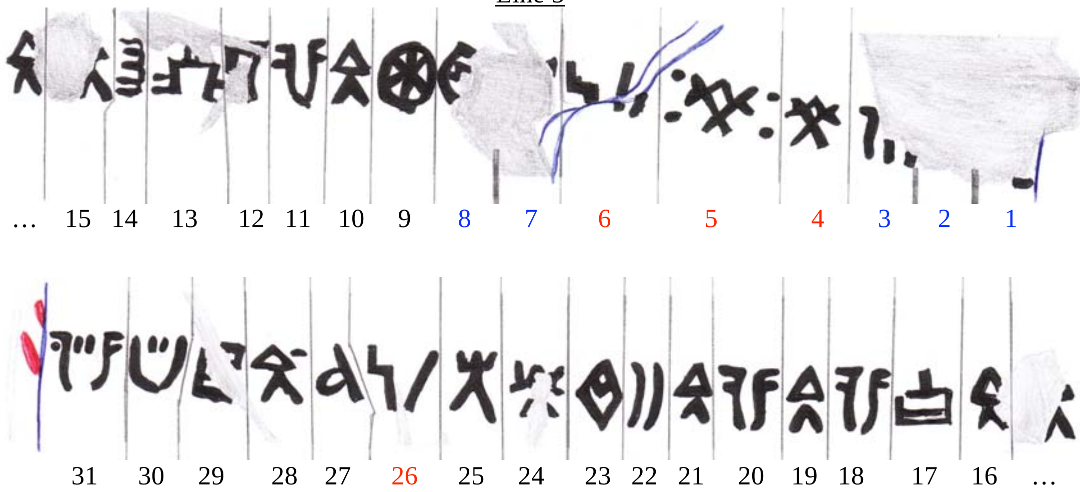
Figure 5. Segmentation of text.



Line 4



Line 5



Line 6



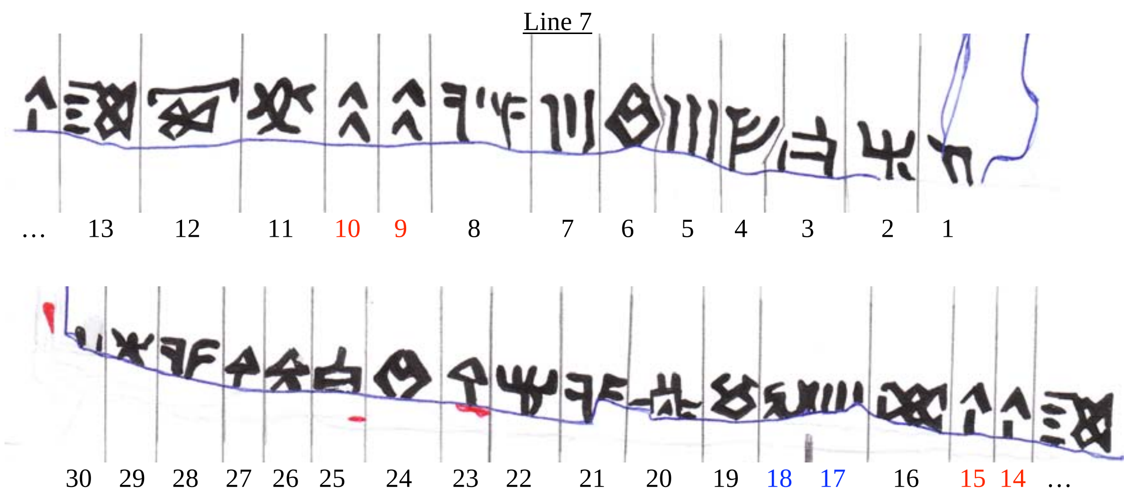


Figure 6. Alternative segmentations.

Position	Chosen segmentation	Alternative no. 1	Alternative no. 2	Alternative no. 3
1.18				
1.19, 2.7, 2.16, 3.16, 3.21, 4.17, 6.14 and 7.7				
1.24-25				
1.31-32				
2.3-4				
3.11-12				
5.4-5				

5.6				
5.26				
6.15-16				
6.18-19				
6.31				
7.9-10				
L7.14-15				

Figure 7. Kabul sign list with allographs.

1					2						3		
	1.32	2.4	2.18	1.15		3.32	1.21	6.16	3.14	2.3		2.20	1.28

4							5			
	4.20	1.13	5.22	1.26	4.15	6.15		2.6	1.38	4.12	7.5

6		7		8		9				10		
	2.25		6.31		1.10		5.26	5.6	1.31		7.15	7.14

11				12								
	7.27	7.23	2.5		5.19	7.26	5.16	3.6	4.16	6.24	5.21	4.19

	13							14		
5.10		1.23	3.4	3.26	1.8	5.28	6.32		7.5	7.10

		15		16			17			
4.6	4.13		1.4		6.22	1.6		3.3	2.18	6.19

18				19		20		21		
	7.12	2.10	1.3		7.16		7.13		1.36	1.17

22						23					
	6.13	5.17	7.3	7.25	4.35		3.30	6.23	3.11	3.12	6.2

24				25			26				
	3.25	3.28	1.20		1.22	1.37		4.18	5.23	4.24	7.6

				27			28		
7.24	4.36	4.21	3.19		6.5	1.1		1.27	1.29

29			30		31			32		
	2.19	1.35		7.19		5.13	6.7		2.30	7.20

33		34		35		36				
	2.15		1.5		1.18		1.19	2.7	2.16	6.14

			37						
3.21	7.7	3.27		5.11	2.17	3.19	4.14	2.8	1.9

								38	
5.18	5.20	1.30	7.21	3.10	3.17	7.28	4.34		5.31

		39			40				41		
7.8	3.5		5.30	1.34		6.17	6.28	1.7		3.22	5.27

	42					43		44		
6.21		1.33	5.14	6.30	1.2		7.22		4.23	7.4

45						46		47	
	5.9	6.3	1.16	6.12	3.31		6.4		5.14

48							49		50	
	5.25	4.33	4.30	3.9	2.9	7.29		4.29		4.2

51		52		53		54			
	3.20		2.2		6.9		1.12	1.25	4.26

55		56					57	
	4.7		5.4	6.8	7.11	1.11		4.25

58		59		60		61		62	
	5.5		1.24		6.20		6.29		4.28

Figure 8. Near identical graphemes in the Kabul manuscript and the Indus script (Parpola 2000: 70-78).

Kabul grapheme	Indus grapheme
1	125a
2	147a
3	127a
4	148a
5	150a
6	131g
8	145a
10	142a
11	189a
14	137a
21	278g
25	341f
26	376f
27	273a
30	233d
32	288c
36	302bb
37	311j
38	313d
39	297a
43	97l
45	368e
53	131 (Wells 1998)
54	60a
55	70c
56	72d
58	73a

Figure 9. Near identical correspondences in mirror image (Parpola 2000: 70-78).

Kabul grapheme	Indus grapheme
17	241e
18	242d
19	229b
40	162b
41	167a
42	107m
44	91aa
47	354k
51	27e

Figure 10. Strong resemblances with fewer strokes (Parpola 2000: 70-78).

Kabul grapheme	Indus grapheme
13	9n
15	192l
22	277h
23	272p
24	273s
34	197c
48	7h
52	4i

Figure 11. Uncertain resemblances due to obstructions (Parpola 2000: 70-78).

Kabul grapheme	Indus grapheme	
33	307h	305a

35	319w

46	360a	361h

50	1q

Figure 12. Fair resemblances (Parpola 2000: 70-78).

Kabul grapheme	Indus grapheme
16	210f

7	131d

28	352

60	83f	81d	116l

Figure 13. Modifications of other Kabul signs.

Modified Kabul grapheme	Similar Kabul grapheme
20	19

29	25

31	32

59	58	55

Modified Kabul grapheme	Similar Kabul grapheme	Similar Indus grapheme
49	48	6a

57	56	64b

Modified Kabul grapheme	Similar Kabul grapheme	Indus symbol
12	13	Krishna Rao 1998: Figure 24

Figure 14. No obvious resemblances.

9	61	62

Figure 15. Comparison of alternative segmentations (Parpola 2000: 70-78).

1.19, 2.7, 2.16, 3.16, 3.21, 4.17, 6.14 and 7.7	126a

1.32	5.6	5.26	Rhd-138 A (Shah and Parpola 1991: 365)	Rhd-133 A (Shah and Parpola 1991: 365)

2.3-4	126a

3.11-12	510 (Wells 1998: 69)

6.31	175d

L7.14-15	243 (Wells 1998: 64)

Figure 16. Comparison of alternative allograph divisions (Parpola 2000: 70-78).

Allographs of Kabul grapheme 23		
	6.23	3.11
Two distinct signs according to Parpola 2000		
	272p	292j

Allographs of Kabul grapheme 41		
	3.22	5.27
Two distinct signs according to Parpola		
	167a	214d

Allographs of Kabul grapheme 26		
	4.21	3.19
Same grapheme according to Parpola		
	376f	376d

Allographs of Kabul grapheme 42		
	6.30	1.2
Same grapheme according to Parpola		
	107m	107t

Allographs of Kabul grapheme 44		
	4.23	7.4
Same grapheme according to Parpola		
	91aa	91bb

Figure 17. Most common Indus graphemes (Parpola 2000: 70-78 and Wells 1998: 60-70).

rank	Indus sign	Total occurrences
1	311	855
2	127	481
3	91	242
4	376	232
5	150	215
6	296	194
7	60	187
8	72	179
9	189	157
10	107	151
11	368	137
12	354	131
13	70	127
14	302	123
15	126	120
16	68	116
17	148	115
18	278	99
19	125	82
20	272	76
21	312	75
22	101	70
23	87	66
24	319	66
25	162	65
26	88	65
27	192	64
28	7	62
29	4	58
30	273	57
31	66	55
32	337	51
33	9	49
34	13	47
35	83	44
36	214	43
37	147	42
38	260	40
39	97	38

40	134	38
41	191	38
42	204	37
43	361	36
44	1	35
45	156	35
46	131	34
47	198	33

48	225	33
49	145	30
50	54	27
51	187	27
52	289	27
53	307	27
54	283	26
55	277	25

56	177	24
57	273	24
58	288	24
59	48	23
60	313	22

Figure 18. Comparison of the number of Kabul sings and the number of Easter Island sings that appear in the top 60 most common Indus signs, given in %.

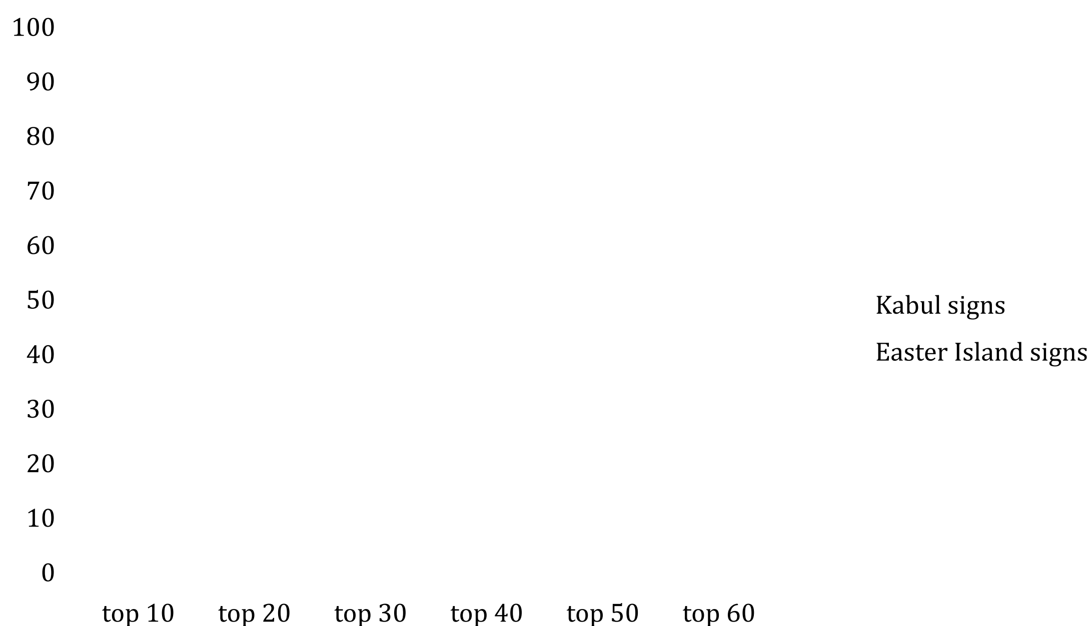


Figure 19. Number of most common Kabul sings which are amongst the top 60 Indus signs, given in %.

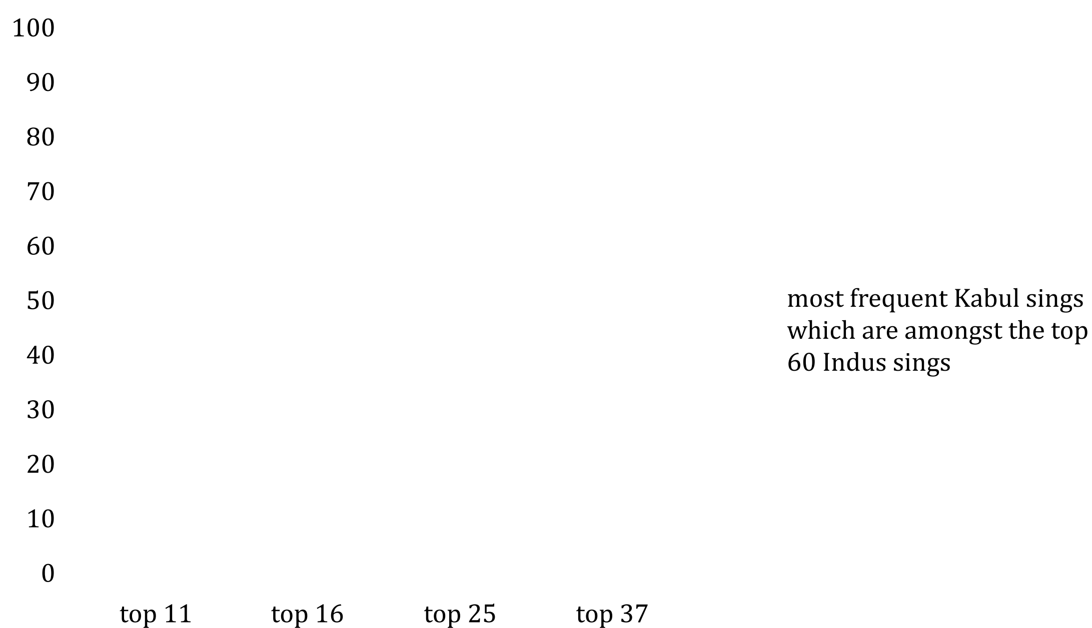


Figure 20. Most common Kabul graphemes.

Rank	Kabul grapheme	Total occurrence in the Kabul manuscript	In the top _ of the Indus graphemes
1	37	14	10
2	12	9	-
3	26	8	10
4	36	7	20
5	4	6	20
6	13	6	30
7	48	6	30
8	2	5	40
9	22	5	60
10	23	5	20
11	45	5	20
12	1	4	20
13	5	4	10
14	14	4	Lower than 60

15	42	4	10
16	56	4	10
17	9	3	-
18	11	3	10
19	17	3	Lower than 60
20	18	3	Lower than 60
21	24	3	30
22	38	3	Lower than 60
23	40	3	30
24	41	3	Lower than 60
25	54	3	10
26	3	2	10
27	10	2	Lower than 60
28	16	2	-
29	21	2	20
30	25	2	Lower than 60

31	27	2	60
32	28	2	Lower than 60
33	29	2	-
34	31	2	-
35	32	2	60
36	39	2	Lower than 60
37	44	2	10